DOCUMENT RESUME

ED 277 241 FL 016 170

AUTHOR Lofstrom, Jonas
TITLE Reports on Swedish Projects Relevant to the Study of Impromptu Speech: Sprakdata.
PUB DATE 82
NOTE 8p.; In: Enkvist, Nils Erik, Ed. Impromptu Speech: A Symposium. Papers Contributed to a Symposium on Problems in the Linguistic Study of Impromptu Speech (Abo, Finland, November 20-22, 1981); see FL 016 148.
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classification; *Communication (Thought Transfer); Computer Oriented Programs; Correlation; Discourse Analysis; Foreign Countries; Interaction; Language Patterns; Language Research; Native Speakers; *Oral Language; *Research Methodology; Research Projects; *Semantics; Sociocultural Patterns; Speech Acts; Spontaneous Behavior; Standard Spoken Usage; *Swedish; Uncommonly Taught Languages; *Verbal Communication
IDENTIFIERS *Spontaneous Speech

ABSTRACT

Sociolinguistic Studies in a Data Base of Spoken Swedish (Sprakdata) is a project to describe the language from a semantic-communicative point of view. The purposes are to (1) investigate the possibility of grouping informants based on results of a semantic-communicative analysis of recorded spoken language and (2) examine possible interrelationships between these groups and social data. The language corpus consists of about 100 recordings of free conversation, comprising about 500,000 words. Detailed linguistic criteria concerning how a given message is communicated have been established for ten semantic-communicative categories. The categories emphasize modality and causal relations, attitudes toward the message receiver, and the form of the message. The corpus has been transcribed and coded for machine reading so that the linguistic analysis can be done interactively at display terminals. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277241

REPORTS ON SWEDISH PROJECTS RELEVANT TO THE STUDY OF IMPROMPTU SPEECH

III. SPRAKDATA

Jonas Lofstrom
Goteborgs universitet

FL 016 170

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Enkvist

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

III. Språkdata

Jonas Löfström
Göteborgs universitet

The project "Språksociala studier i en talspråklig databas"[1] aims at describing spoken language from a semantic-communicative point of view. The purpose of the project is (1) to investigate the possibility of grouping informants on the basis of the result of a linguistic (semantic-communicative) analysis of recorded spoken language, and (2) to establish possible interrelations between these groups and social data. Theoretical considerations and practical work have led to the conclusion that, for the purpose of the project, free conversation is the optimal speech situation. A random selection of informants have been asked to choose a partner for free conversation, to be recorded by the informants themselves without the assistance of project personnel. A corpus of about 100 recordings, comprising some 500,000 running words, has been gathered.

Detailed linguistic criteria distinguishing about ten semantic-communicative categories have been formulated. The linguistic analysis largely concerns the question how a given message is communicated. Among the categories thus studied, which are both illocutionary and propositional in nature, emphasis has, in particular, been laid on modality and causal relations as well as attitudes toward the receiver and the form of the message.

From a theoretical point of view, a speech situation where both participants can be said to play their habitual roles in an appropriate setting, talking about well-known subjects of their own choice, yields the best conditions for gathering a corpus of casual speech. That is to say that as few factors as possible should make the speakers inclined to reflect on their language (and

[1] I.e. "Sociolinguistic Studies in a Data Base of Spoken Swedish". The project is financed by grants from SÖ (the National Board of Education) and HSFR (the Swedish Humanistic and Social Science Research Council) and is carried out at the Department of Computational Linguistics (Språkdata), University of Göteborg. Leader of the project is Sture Allén.

possibly use a planned or high status language), and as many factors as possible should favour fluent, natural speech.

The methodological consequences of the aim to achieve natural role relations for the recordings were primarily to let the participants do their recordings of free conversations themselves, without the assistance of any project personnel. We also preferred the recordings to take place at home or at work, i.e. at a well-known place with a partner of their own choice. It turned out that the majority of the informants had access to a cassette recorder. (The others used one of the project recorders.) This also means that they were used to the technical devices needed. These are circumstances which favour our endeavour. The conversations turn out, actually, in many cases, to be highly improvised. The speakers, knowing each other well, permit themselves to associate freely and change the subject accordingly. Some of the recordings are poor from a technical point of view (that is the risk one has to accept for not involving technical personnel) or characterized by the nervousness of the speakers. Most of the recordings, however, are astonishingly natural.

From the point of view of the theme Impromptu Speech the corpus thus provides a relevant body of material. The project, focusing on a semantic-communicative analysis, does not directly study the impromptu aspect of the conversations, but taking the content as a starting point for the analysis one has to deal with different forms of expression. In the conversations some of these, of course, are characterized by the conditions of casual speech. A few examples of observations of these phenomena in the corpus will be given below.

What, then, constitutes impromptu speech in these conversations? (Without giving a definition of impromptu speech we here focus on such cases of improvisation that the speech situation in some way has brought about.) To answer the question we need to take into consideration the conditions of a conversation. The physical environment being common to the two participants, a lot of information need not be conveyed explicitly. The speech situation involves interaction which requires collaboration between the interlocutors. They thus have to make sure their partner follows, and, on the other hand, adapt themselves to the answers and reactions they get. Another aspect of the cooperation is that both speakers have to defend their own position and make sure they can present their opinion. One has to plan this while the interlocutor is talking and maybe make one's statement in close connec-

tion with his speech. Neither of them knows how the dialogue will end. The aim of the conversation is normally not fixed, but there is room for free associations.

Another important condition of spoken language is that the different pieces of information have to be presented in a temporal sequence. The speaker's execution of his talk often tends to begin with a general picture. The important information, modifications and appeals do not appear until the end of the utterance.

From the point of view of the single speaker the spatio-temporal situation that the conversation constitutes also includes the speech of the interlocutor. This means he will get into situations where he has to complete someone else's utterance, which involves a component of improvisation. Linguistically this means that the line of one speaker does not necessarily include all the component parts of a sentence that would correspond to the complete sentence. Some of them are left out (the context is clear enough), it would be inappropriate to be redundant, others are said by the interlocutor. In a written language perspective this would be characterized as ellipsis. This can be final or intitial depending on the context.

In the dialogue one constantly has to check whether one's interlocutor follows. Short lines with appeals like *don't you think, isn't it, I guess* give you the opportunity to stop and explain, if you don't get positive feedback. In your capacity as the listener you can make sure you correctly understood what the speaker said by adding a phrase or a clause in the form of a question to his sentence:

A: men jag har sett det förr
(but I have seen it before)

B: när du har varit utomlands e l l e r
(when you have been abroad, you mean
(literally: 'or')).

The cooperation does not give you any time for preparation, you must answer directly or adjust your line to what has just been said. On a linguistic level this is to say that you work out your line with an expression or construction of your interlocutor as the syntactic starting-point.

In free conversations the interlocutors easily change the subject by way of association. In order to express his idea, the speaker must signal the change

of subject by indicating what area he is interested in. The spontaneous change, maybe caused by an idea several steps back in the conversation, is not necessarily indicated by linguistic means. But he can refer to a word or idea in the preceding speech, or he can give a general indication like *by the way...* These conditions characterize a lot of spoken language: there is little time for preparation, the general idea is to be presented first, and the restrictions on short time memory (in the speaker and the listener) do not permit an immediate indication of the logical structure. The ideas have to be presented in the order they appear. As has often been pointed out, spoken language tends to give important information at the end of an utterance, because many circumstances that the speaker does not control influence his planning of the speech.

The analysis of content categories in the project shows how the use of one category can support the interpretation of another, in other words a kind of redundancy. In the impromptu speech perspective there is another typically interactive phenomenon in the use of content categories. The interlocutors rather often cooperate in constructing (expressing) a token of a content category. This cooperation between the speakers constitutes a difficulty in the semantic analysis of spoken language. On the other hand, it is a major feature in the characterization of conversation and impromptu speech.

In the conversation, the line of thought of the speaker such as it is interpreted by his interlocutor is sometimes completed by the latter. The cause indicated by one spekaer is followed by an expression of the effect from the other speaker.

B: Jag ska bada först *(I am going to have a bath before)*	(interpreted CAUSE)
A: ja då kommer du vid femtiden då *(well, then you'll come by five o'clock)*	*EFFECT*

The listener interprets a cause in the line of the speaker and indicates the effect. The two instances are not completely comparable and therefore kept apart. In the following example D interprets A's explicit effect (1) as containing also a cause (2) for which she states the effect (2):

A:	*kan man böja dom* *(can one bend them)*	
B:	*nej det ska du inte helst inte göra* *(no you've better not do that)*	(interpreted CAUSE 1)
A:	*nej då lägger jag dom här så länge* *(well, then I'll put them here for a while)*	EFFECT 1 (interpreted CAUSE 2)
D:	*då tappar du dom sen* *(but then you'll loose them)*	EFFECT 2
A:	*nej bara så länge* *(no, only for a while)*	

Modality shows a similar pattern. The listener sometimes makes a major contribution to the proposition the modality of which is expressed by the speaker.

B: *så det är riktigt men ändå så så blir det att vi får*
(so that is correct but it still turns out that we
shall have to)

A: *ja klä skott ja*
(yes, suffer from this)

The informants present different patterns both in their interaction and in the frequency of the different categories. One aim of the project is to find out whether the informants can be grouped according to their use of these categories.

The transcription of the corpus of conversations has been made available in machine-readable form by means of encoding. The linguistic analysis is carried out interactively at display terminals and the results are stored in such a way that access is possible both via the content categories and the linguistic expression. This flexibility makes it possible to compute the results in various ways. There is also a possibility of relating interactionally connected entities which are of relevance for the analysis of impromptu speech.

REFERENCES

Abelin, Å., Gadelii, E. & Löfström, J.: Om modalitet i talspråk. (On modality in spoken language.) Rapporter från Språkdata 16. Göteborgs universitet. 1981.

Löfström, J.: Konkordanser till lingvistiskt bearbetad text. (Concordances to Linguistically Prepared Text.) In: Föredrag från 2:a svenska kollokviet i språklig databehandling i Lund 1979. Eds. C. Thavenius och B. Oreström. Survey of Spoken English. Lunds universitet. Lund 1979.

——— Innehållskategorier i samtal i Göteborg. (Content Categories in Conversations from Göteborg.) To appear in Förhandlingar vid sammankomst för att dryfta frågor rörande svenskans beskrivning 13. Helsingfors universitet.

Löfström, J., & Norén, K.: Teori- och metoddiskussion kring projektet Språksociala studier i en talspråklig databas. (A Discussion of Theory and Method in the Project Sociolinguistic Studies in a Data Base of Spoken Swedish.) Rapporter från Språkdata 5. Göteborgs universitet. 1977.

Löfström, J., & Palmgren, K.: Om innehållskategorier i talspråk – exemplet orsak-verkan. (On Content Categories in Spoken Language – the Case Cause-Effect.) Förhandlingar vid sammankomst för att dryfta frågor rörande svenskans beskrivning. 11. 1978, p. 125-133. Stockholm 1979.

MacDowall, M.: Transkribera talspråk. Bidrag till en transkriptionsmanual för projektet Språksociala studier i en talspråklig databas. (Transcription of Spoken Language. Contributions to a Manual of Transcription for the Project Sociolinguistic Studies in a Data Base of Spoken Swedish.) Rapporter från Språkdata 4. Göteborgs universitet. 1977.

Palmgren, K., & Löfström, J.: Om språkliga innehållskategorier. (On Linguistic Content Categories.) Rapporter från Språkdata 8. Göteborgs universitet. 1978.

Språksociala studier i en talspråklig databas. Förstudierapport. (Sociolinguistic Studies in a Data Base of Spoken Swedish. Report from a pilot study.) Kerstin Palmgren, Jonas Löfström, Christian Sjögreen och Sture Allén. Stencil. Språkdata. Göteborgs universitet. 1978.

Thelander, M.: Språksociologiska metodfrågor. Kommentar till en projektbeskrivning. (Methodological Problems in Sociolinguistics. Comments on a Project Plan.) Rapporter från Språkdata 2. Göteborgs universitet. 1977.